DOCUMENT RESUME

ED 204 382 TM 810 372

AUTHOR Lennon, Roger T.
TITLE A Time for Faith.
PUB DATE 15 Apr 81
NOTE 11p.: Presidential Address presented at the Annual Meeting of the National Council on Measurement in Education (Los Angeles, CA, April 11-17, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Educational Testing: Evaluation Needs: *Testing Problems

ABSTRACT

The spread of testing to its present near-universal status in American education is ascribed in part to the fact that tests have been found generally superior to suggested alternative ways of obtaining information about human behavior. Where alternative methods are used, it is from necessity rather than choice: there are important human characteristics for which no adequate tests exist. The approximately fifty million Americans taking part in some kind of formal learning experiences give rise to a vast array of information needs, including data about their aptitudes, learning difficulties, achievement status, program effectiveness, etc. These information needs present the measurement profession with challenge and opportunity. It is suggested that measurement practitioners have been over-reacting to critics and insufficiently confident in their own contribution. The acceptance and influence of testing in American education should be a source of pride to them. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
R. T. Lennon
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED204382

A TIME FOR FAITH*

Roger T. Lennon
Harcourt Brace Jovanovich, Inc.

It is a question whether the tradition of having a presidential address follow this annual breakfast is harder on the president or the members. The after-meal speech is a special art form, with its own rules: it must not be too heavy, lest it disturb the flow of gastric juices; it must not be too light, lest the audience feel deprived of a message; it should be neither too bland nor too controversial; and most of all, it should not be very long. Like my predecessors, I have tried within these constraints to craft a few remarks. Like the Gettysburg Address, this speech will be brief; unlike it, it will be one which indeed "the world will little note, nor long remember" -- luckily so, in all likelihood.

Given these situational expectations, I decided not to deliver what I call my "gun-lobby" speech, whose theme is "Tests don't kill people; people kill people." That is no message for nine o'clock in the morning. Rather, I shall take advantage of the occasion to give voice to a few deeply held sentiments about measurement and its practitioners -- and if, in the course of these remarks, some of my prejudices accidentally show through, that is no accident.

TM 810 372

* Presidential Address, National Council on Measurement in Education, Annual Meeting, Los Angeles, California, April 15, 1981.

In a brief sixty or seventy years, educational measurement has moved from a cutting-edge, innovative practice, to its present state as a pervasive, virtually universal feature of American education. The diffusion of standardized testing through American schools has been almost a text-book example of Paul Mort's model of adoption of a successful educational innovation. Where a measurement text in the nineteen-twenties could carry the title "The Objective or New-Type Examination," today it would be hard to find a school district without a testing program of some kind. This diffusion has been accompanied by a continuing increase in scholarly underpinnings, reflected in the appearance of professional organizations such as NCME, specialized journals to report on developments in theory and practice, doctoral programs, governmental support and intervention, and, from the beginning, critical, sometimes hostile attention from lay and professional observers. I have no intention of rehearsing these critical views here; that was done admirably by Jay Millman in his presidential address a year ago. I invite your attention only to the historical fact that over the past sixty years the trend in the use of tests has been steadily -- one is tempted to say inexorably -- upward.

I do not infer from this history that we in measurement have grounds for complacency -- far from it. I would not even argue that the amount of testing, or the growth in it, is ipso facto good (though it has certainly made my life more pleasant). I do suggest that the increase in

testing, in the face of ever more vigorous criticism, is a phenomenon that we in measurement have perhaps not fully appreciated. I had occasion recently to re-read some talks I gave in the late '50's and early '60's -- a discouraging exercise -- and found I was saying in those years that "I had never known testing to be so controversial." Well! Little did I know! But what happened? It was as if test users were oblivious to the critics -- even with respect to the types of tests under most severe attack.

I do infer from the continued flourishing of testing that no satisfactory substitutes for it have appeared. Those who find fault with standardized testing have continually urged the search for alternatives to it. In their conferences and in their publications, they call for reduced dependence on standardized tests in favor of use of alternative ways of gathering information about subjects. And what alternatives do they propose? Time after tired time, we are served the same shopworn list: observation schedules, anecdotal records, work products, interviews, contracts, "whole-child" assessment -- you know the list well. And each time these alternatives are advanced as if our literature were not replete with demonstrations that they compare very poorly indeed with standardized tests by any conventional measures of reliability or validity. To encourage the innocent to root around in the rubble of discredited modes of study of human behavior, in search of some overlooked assess-

ment "jewels," is to dispatch a new band of Argonauts in quest of a non-existent Golden Fleece.

It is as if those who are unhappy with the automobile because it pollutes the atmosphere, or consumes unreasonable amounts of energy, or whatever, ask that we search for alternatives to it as a form of private transportation, and then recommend kiddycars, roller-skates, bicycles, sleighs, or even call for the return of the horse. What kind of nonsense is this? There are legitimate functions for kiddycars and bicycles; there are legitimate functions for interviews, behavior ratings and the like, but they are not substitutes for tests. We resort to them out of necessity, not by choice, in situations where adequate tests are not available. The simple truth is that there has not been, is not now, and will not be any device or scheme which is as efficient as the well-crafted test for capturing relevant samples of human behavior, rendering them in a form that permits analysis, manipulation, and interpretation, and permitting their evaluation as elements in inferential chains leading to diagnosis, understanding, and prediction.

And do you know something? Deep down, everyone -- even our harshest critics -- knows this, for characteristically their diatribes conclude by saying, "but, of course, we have to have tests." The American public knows it, if we may judge from the findings of recent Gallup polls on attitudes towards testing, and teachers know it, if we may credit recent surveys of their opinions.

Nor am I greatly moved by the torrents of crocodile tears shed for measurement's supposed neglect of qualities such as creativity, artistic ability, leadership, industry, honesty, and other eminently desirable human qualities. They are qualities that have proven highly intractable to measurement approaches, not necessarily because of limitations inherent in those approaches, but for more fundamental reasons having to do with definitional and value considerations. If these characteristics are not so well represented in the measurement armamentarium as are cognitive abilities, it is not because the measurement fraternity accords them less honor than cognitive qualities, nor because of failure to attempt measurement in these domains. Rather it is to the credit of measurement people that through their efforts the difficulties and the limitations of attempted measurement in these domains have been made known and the public spared from unwitting reliance on inadequate measures.

There are fifty million people across the country engaged in some type of institutionalized learning activities. They range from pre-schoolers through school-age children, teenagers, college students, and adults. All of them want and need to know how well they are attaining the goals set for them by society or by themselves. Their institutions need to be in a position to assess the effectiveness of the learning experiences which they provide, to modify them where appropriate, and to allocate resources to

yield the greatest benefits. A lot of these learners have difficulties; we should be trying to help them understand and cope with their learning problems. We have only lately begun to devote adequate attention to the needs of the handicapped. Their afflictions, whether of mind or body or spirit, cry out to us for deepened understanding, assessment and intervention. Society would benefit from an inventory of the talents of these learners, and greatly needs information about their progress, for a variety of types of social planning. The movement of these learners through the educational system is clearly facilitated by the availability of relevant information about them.

This list of information needs -- and it includes only the most obvious from a long, long list -- is nothing other than the agenda for the measurement profession. There are human needs to be met, and we can help uniquely, as we have been helping, to meet them with our special brand of expertise. If there are those who find it more fun to stand on the sidelines, pick their nits, and throw spitballs at the players, than to be out on the field where the action is, that is their problem. Our problem is to be about the sharpening of our instruments, the deepening of our understanding of the processes and products that we attempt to measure, the improvement of the applications and uses of test data, and the sharing of insights with our colleagues and students -- precisely the kinds of things we do in meetings such as this. And our future will depend, as has

our past, on how well we achieve these goals, not on how many debaters' points we score against our critics.

If physicians and surgeons were to withdraw to the laboratory or invoke a moratorium on the rendering of services, pending the time when their diagnoses, their medications, and their treatments achieved a state of less fallibility than now prevails, the health of all would surely be the poorer for it. If we, either in an excess of reaction to critical voices, or under the demands of ill-conceived legislation, abandon or diminish our role in education, then education will surely be the poorer.

NCME members, like their counterparts in Division 5 of APA, Divisions D and H of AERA, and AMEG, share common values -- objectivity, open-mindedness, dedication to evidence, judiciousness. We honor, and not merely with lip service, caring for persons, social justice, equality of opportunity, and equity. Curiously, it is precisely because we have these concerns that we tend to react to criticism and attack on our activities, and even more our motivations, with puzzlement, confusion, trepidation, and distress. We are accustomed to carrying on our affairs according to certain canons of scholarly discourse -- not in a strident, politicized, adversarial context. We are flattered to be invited to "debate" on T. V., and then chagrined to discover that we have been had by persons whose mission is entertainment, not enlightenment. We listen in bewilderment to persons who accuse us of striving to foster and perpetuate discrimination. We hear judicial pronouncements to the

effect that, yes, our tests are biased, and are guilt-ridden. In our naivete, we even treat with earnestness and courtesy the strictures of those whose only goal is our destruction. We tend to say, almost plaintively, in Tom Fitzgibbon's expression of a few years ago, "Why don't they love us any more?" -- the while presenting to the world an unseemly, and inappropriate, image of disarray, or at least uncertainty.

Testing should not expect or want to be immune from the hurly-burly, catch-as-catch-can contention that is the heart of the democratic process; but it must not fall into the trap of allowing its enemies to establish the terms of discourse or the conditions of debate. I am not suggesting that we forsake any of our traditional scholarly values, but I think it is time for us to reassert a fundamental faith in the worth of what we are doing, such as animated Alfred Binet, Lewis Terman, E. L. Thorndike, Truman Kelley, E. F. Lindquist, and all those others in whose footsteps we have elected to follow. In such affirmation, there need be no implication that all is well in measurement's house; we are neither blind nor foolish. If we reject Cassandra with her gloomy reading of measurement's role and future, we need not embrace Pollyanna and her euphoric delusions about measurement. No one who reads our publications or attends our meetings can suppose that we are lacking in self-criticism. That self-criticism, and the devotion of scholars to the improvement of both the theory and the practice of our craft, are evident all about us.

We should remember that ours is a young technology. I have no doubt that fifty years from now, some of our practices will strike scholars of that time as quaint, ill-considered, or even hurtful -- maybe even deserving of some of the opprobrium now heaped on them. But today there are these information needs to be met, and I think we have the obligation both to do that job as well as the present state of our art permits, and to say to critics, we will listen to you when you have something demonstrably better to offer.

I do not believe that it is coincidental that testing has taken root and flourished in the United States to a degree that has not occurred elsewhere around the world. Clearly here the climate was right and the soil was ready for the flowering of this child of the psychological laboratory and the so-called scientific movement in education. I am not enough of an educational historian or sociologist even to suggest hypotheses as to why it has been that measurement has been so well received in our society. Yet there are aspects of American education in the twentieth century that seem to me predisposing toward, if not indeed requiring, the kind of standardized testing to which we have become accustomed. The circumstance that education in America has been a state concern, and not a national concern as in most other nations around the world, has enhanced the utility of the nationally-normed test as a source of some common currency in which to express the results of schooling across states as well

as within states. But at a deeper level, the commitment of our society to universal education and, in the latter half of this century, to repairing the ravages of the damage done to minority groups in the educational as well as other realms, has placed a premium on the kinds of information about learners and about the efficacy of educational programs that tests are particularly well suited to provide. Tests, moreover, seem to me especially responsive to a deep pragmatic impulse in the American spirit: they provide workable mechanisms for managing the colossal task of universal education, for guiding subjects through the system into its various levels and components, for monitoring the learning process, and thus, in the long run, providing direction for altering instructional programs. They do this even though everyone knows they are far from perfect; but almost everyone who has to make decisions feels they are far more helpful than hurtful.

If associating the merits of measurement with special characteristics of the American scene seems to be a bit of flag-waving, I plead guilty. One of the small rewards of advancing years is that one may be corny without feeling embarrassed. And as I am happy to be an American, I have been proud to play a part in that effort to bring the fruits of measurement scholarship to bear on human needs which is test publishing. I invite you, my colleagues, to share this pride in our common work: we are engaged in a worthy endeavor.

####